DOCUMENT RESUME

ED 252 940 EA 017 450

TITLE	Competency Assessment and Testing. Minimum Standards Implementation Series.
INSTITUTION	Ohio State Dept. of Education, Columbus. Div. of Elementary and Secondary Education.
PUB DATE	83
NOTE	77p.; For related documents, see EA 017 445-459.
PUB TYPE	Guides - Non-Classroom Use (055)
EDRS PRICE	MF01/PC04 Plus Postage.
DESCRIPTORS	Academic Standards; *Competency Based Education; Elementary Secondary Education; Evaluation Criteria; Mathematics Achievement; Minimum Competencies; *Minimum Competency Testing; Reading Tests; *State Standards; Student Certification; *Testing Programs; Test Reliability; Test Reviews; *Test Selection; Test Use; Writing Evaluation
IDENTIFIERS	*Ohio; Ohio State Board of Education

ABSTRACT

This is one of three publications prepared by the Ohio State Board of Education to assist schools in implementing a competency based education program, as required by the 1983 revised minimum standards for elementary and secondary schools. This document addresses the assessment requirement, with emphasis on defining, comparing, and contrasting various options for meeting this requirement. The first section provides an overview of the competency standard, lists the steps to follow in establishing a competency program, outlines decisions that a district may or may not make, describes the various uses of assessment/test results, and identifies important issues related to the testing program. The second section examines each of these highlighted issues more closely, discussing several options for each. In the third section, other important issues are identified, procedures for dealing with them are detailed, and implications for a district are discussed. These issues include staffing, resources, cost, assuring fairness, technical considerations, evaluation/review, and coordination with existing programs. The fourth section provides suggestions for developing a management plan that can be used to monitor progress. In the final section, sample timelines for activities illustrate the sequence in which events may occur. A bibliography is included. (TE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED252940

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

[ ] Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

F. B. Walter

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

EA 017 450

IMPLEMENTATION SERIES

STATE BOARD OF EDUCATION

Richard C. Glowacki, President	Toledo
Wayne E. Shaffer, Vice President	Bryan
Jean E. Bender	Akron
Wallace E. Blake	Zanesville
Paul Brickner	Willoughby
Jack C. Hunter	Youngstown
Virginia E. Jacobs	Lima
Robert H. Johnson	Springfield
Mary R. Lindner	Cincinnati
William E. Moore	Woodsfield
Sue Ann Norton	Westerville
Paul F. Pfeiffer	Massillon
Edwin C. Price, Jr.	Cincinnati
Chester A. Roush	Kettering
Anthony J. Russo	Mayfield Village
Lucy Russo	Cleveland
Steven Schecter	Cleveland Heights
Patricia Smith	Worthington
Sally R. Southard	Oxford
Jo Thatcher	Portsmouth
Karen L. Worley	Elyria

Franklin B. Walter
Superintendent of Public Instruction

MINIMUM STANDARDS IMPLEMENTATION SERIES

STEERING COMMITTEE

G. Robert Bowers, CHAIR
Assistant Superintendent

Irene G. Bandy
Assistant Superintendent

Nancy Ann Eberhart
Associate Director
Office of Inservice Education

Robert W. Evans
Assistant Superintendent

James W. Miller
Director
Division of Educational Services

Mary J. Poston
Director
Division of Elementary and Secondary Education

COMPETENCY ASSESSMENT AND TESTING DEVELOPMENT TEAM

E. Roger Trent, CHAIR
Consultant
Division of Educational Services

Randall Flora
Consultant
Division of Elementary and Secondary Education

Margaret B. Kasten
Consultant
Division of Elementary and Secondary Education

Hazeldean Stent
Consultant
Division of Elementary and Secondary Education

BEST COPY AVAILABLE

FOREWORD

With the adoption of the 1983 minimum standards, the State Board of Education has established conditions to promote a general education of high quality in all chartered schools. The State Board of Education recognizes its responsibility to give direction to the system of education in the state. At the same time, there is flexibility within standards to establish educational programs that are responsive to local needs.

Minimum Standards for Elementary and Secondary Schools 1983 provide the potential for greatly improving education in Ohio. That potential can be reached through well-structured local and state leadership. To assist responsible implementation of minimum standards, staff in the Ohio Department of Education have prepared an implementation series consisting of fourteen documents. A distinction is made among the publications. Only one, *Procedures for School Evaluations*, is a compliance document. It contains the standards in a format readily used for self-evaluation. School personnel will find information about what is needed for a district and schools within the district to comply with standards. In the remaining publications, an effort is made to specify what the standard requires. Additional information, guidelines, procedural steps, and sample forms may go beyond what the standard requires. The purpose is to assist the local planning process, not to prescribe what shall be in every school in Ohio.

Minimum standards address the public's increasing expectations for schools with a requirement for competency based education in English composition, mathematics, and reading. Three of the publications, *Process Model for Course of Study, Competency Assessment and Testing*, and *Intervention* are aligned in addressing competency based education. A number of options are described for local consideration.

Appreciation is extended to school personnel who contributed to these publications by sharing their time, expertise, and materials. Appreciation is also extended to staff members who worked untold hours to make these publications possible.

Franklin B. Walter

Franklin B. Walter
Superintendent of Public Instruction

TABLE OF CONTENTS

Section		Page
1	INTRODUCTION	4
2	DECISION MAKING	13
	Which Competencies to Test	13
	Assessment/Test Distinction	16
	Which Test to Use	18
	Criteria for Selecting/Constructing Instruments	22
	Special Considerations: English Composition	26
	Special Considerations: Reading	31
	Special Considerations: Mathematics	32
	Assessing Mathematics Problem Solving	33
	When to Test	35
	Guidelines for Use of Results	37
	Communication of the Results	44
	Setting Performance Standards	45
	Providing Intervention	48
	Guidelines for the Handicapped	49
3	COMPETENCY TESTING: OTHER IMPORTANT ISSUES	52
	Staffing	52
	Resources	52
	Costs	53
	Assuring Fairness	53
	Technical Considerations	54
	Evaluation/Review	55
	Coordination with Existing Programs	56
4	GETTING STARTED	57
	Conduct Study	58
	Develop Management Plan	61
	Models for Organizing the Project	62
5	IMPLEMENTATION TIMELINE	63
6	SELECTED BIBLIOGRAPHY	68

PREFACE

Process Model for Course of Study, Competency Assessment and Testing, and *Intervention* are three publications that will be helpful in implementing a competency based education program as required by the Minimum Standards for Elementary and Secondary Schools 1983. Each of the three documents addresses different aspects of effective competency based education programs. Although each book is designed to stand alone, all three are needed for full implementation. The following descriptions show which document should be used for each phase.

During the early planning stages, Sections One, Four and Five of the *Competency Assessment and Testing* document will be most helpful. Included are a list of issues and decisions (Section One), suggestions for how to get started (Section Four), and sample timelines (Section Five).

For suggestions regarding the format for each course of study, see the *Process Model for Course of Study* document Chapters One through Three. Chapter Four of that document deals with the development of pupil performance objectives for the competency programs in English composition, mathematics and reading.

Many assessment and testing decisions are discussed in Section Two of the *Competency Assessment and Testing* document. For each decision several options are evaluated in terms of advantages and disadvantages. In addition, illustrative guidelines for the use of assessment results are offered and considerations related to handicapped students are discussed.

For assistance in planning and implementing an effective instruction/intervention component, see the *Intervention* document. Suggestions regarding the format and design of the intervention component of the competency program are offered in Parts I and II. The elements of effective instructional programs in the three areas for which competency programs are required are discussed in Parts III, IV and V. In addition, many alternative approaches to intervention in the three areas are detailed and illustrative teaching strategies are offered.

Suggestions for planning a review and evaluating the impact of the competency programs are offered in Section Three of the *Competency Assessment and Testing* document.

INTRODUCTION

Section 1
INTRODUCTION

The Minimum Standards for Elementary and Secondary Schools 1983, adopted by the State Board of Education, require that school districts develop and implement competency based education programs. Although the competency based education standard detailed below has far-reaching implications, this document will address mainly the assessment requirement with particular emphasis on defining, comparing and contrasting various options for meeting this requirement. Other documents will address in more detail the implications for curriculum, instruction and intervention.

This first section provides an overview of the competency standard, lists the steps which a district may follow in establishing a competency program, outlines the decisions which a district may or may not make, describes the various uses of assessment/test results, and identifies some important issues to consider related to the testing program. In the second section, each of the decisions highlighted in the first section is inspected more closely. For each decision, several options are discussed. For each option, some advantages, disadvantages and other considerations are explored. In the third section, some important issues are identified, procedures for dealing with some of these issues are detailed, and implications for a district are discussed.

In the fourth section, suggestions are offered for conducting a study to determine what needs to be done and for developing a management plan which can be used to monitor progress.

In the final section, sample timelines for activities illustrate the sequence in which events may occur.

INTRODUCTION

WHAT DOES THE STANDARD SAY?

Section 3301-35-02 (B)(2)
Locally developed competency based education programs shall be implemented for English composition, mathematics, and reading.

- Pupil performance objectives shall be established for English composition, mathematics, and reading.
- Provisions shall be made for periodic assessment of pupil performance, including testing at least once in grades one through four, grades five through eight, and grades nine through eleven.
- Guidelines shall be established for the use of assessment results for instruction, evaluation, intervention, guidance, and promotion decisions.
- Intervention shall be provided according to pupil needs.
- Written policies and procedures shall be adopted by the board of education regarding the participation of handicapped pupils in locally developed competency based education programs. Provisions may be made for exemption of individual handicapped pupils from the requirements of the locally developed competency based education programs.
- Implementation of competency based education in English composition, mathematics, and reading shall begin no later than the 1984-85 school year, with full implementation to be completed no later than the 1989-90 school year.
- Competency based education programs shall be reviewed and updated at least once every five years.

INTRODUCTION

WHAT IS NECESSARY IN ESTABLISHING A COMPETENCY PROGRAM?

The process suggested throughout this publication is based upon the following assumptions:

- Effective competency based education programs are broader than a simple testing system which sorts youngsters into categories. They are instructional and curricular activities which utilize measurement to assess student strengths and weaknesses so that proper corrections or supportive plans can be formulated.
- Assessment is broader than paper and pencil tests. Effective student assessment relies upon multiple methods of measuring student achievement.
- The important aspects of the education process are what happens before measurement and what occurs afterwards as the follow-up to identified strengths or weaknesses.
- Educational programs function best when the initiative, planning and management are kept close to the people. The state has a role to play in charting directions, but the management and day-to-day decisions work best when they are in the hands of a locally elected board of education and locally selected school administrators.
- Education is the shared responsibility of the parent, child, community and school. Each has a vital and fundamental role to play and each must complement the others.

DEFINITION: Although many different definitions of the term "competency based education" are available, in this document the definition contains at least four elements: a list of carefully defined pupil performance objectives, assessments of student performance on those objectives, standards to judge the adequacy of performance, and intervention geared to meet student needs as shown by the assessment results.

The following steps are necessary in establishing a competency program:

- Pupil performance objectives (student outcomes) which are derived from the course of study will be established K-12 for English composition, reading and mathematics.
- Instruction appropriate for helping students achieve those outcomes will be provided.
- Appropriate assessment techniques will be used periodically to monitor each student's performance on the objectives established in three areas.

INTRODUCTION

- The assessment techniques must include testing at least three times (once in grades one through four, once in grades five through eight, and once in grades nine through eleven). These tests will be related to the established pupil performance objectives.

DEFINITION: Unless otherwise noted, testing is used throughout this document to include writing samples and mathematical problem solving, as well as the typical paper and pencil tests which are administered districtwide in at least three grade levels.

- Guidelines will be established for using the assessment results (including testing) in instruction, evaluation, intervention, guidance and promotion decisions.
- Following assessment, intervention based upon student needs will be provided.

DEFINITION: Intervention means alternative or supplemental action designed to remediate, reinforce or support pupil learning relative to specified performance objectives.

- Written policies and procedures regarding the participation of handicapped students will be established and adopted by the board of education.
- The competency based education programs will be provided without discrimination on the basis of color, national origin, race or sex.
- The timeline for meeting this standard is as follows: (a) implementation will begin during the 1984-85 school year; (b) implementation will be completed by the 1989-90 school year; and (c) the program will be reviewed and updated at least once every five years.

INTRODUCTION

WHICH DECISIONS ARE LEFT TO EACH DISTRICT?

- Which of the established competencies (pupil performance objectives) will be tested?
- Which type of test or assessment will be used? How to address the special considerations associated with testing in each of the skill areas (i.e., English composition, reading and mathematics)?
- When the tests will be given (i.e., which grade and when during the year)? Whether or not to test in course areas other than those specified or in more than the required number of grades?
- How the assessment and test results will be used (e.g., for intervention, guidance, evaluation)?
- What type of intervention will be provided for students?
- What relationship will be maintained between the tests used at three or more grade levels and the other assessment techniques used in the competency program?
- How handicapped students will be affected by the competency program?

INTRODUCTION

WHAT USES CAN BE MADE OF ASSESSMENT/TEST RESULTS?

The uses of the assessment and test results alluded to throughout this document can be classified generally into three categories: (1) instructional, intervention, and promotion decisions regarding a particular student; (2) guidance decisions regarding a particular student; (3) educational program decisions such as making school policy or conducting curriculum evaluation. While each is a legitimate use of assessment/test information, the considerations associated with each may be different. The following may help delineate some of these differences.

1. For instructional/intervention/promotion decisions:
   - Assessments may be diagnostic or clinical with a continuing relationship between the student and person making the assessment.
   - Assessments may be made frequently—often each assessment instrument covers a relatively narrow range of performance indicators.
   - Assessment results collected at one point in time may be different from those collected from the same student at another point in time (e.g., results before and after instruction). NOTE: A series of instructional decisions based upon assessment may culminate in a promotion decision.
   - A norm-referenced score (e.g., a national percentile rank on a standardized reading test) is usually not very helpful in making such decisions. On the other hand, a student item analysis or raw score summary for each category of skills covered by that same test may be helpful if those items measure skills which are being taught in the classroom.
2. For guidance decisions:
   - Assessment may be conducted for the purpose of identifying areas of strength and weakness across a broad range of indicators.
   - Although the assessment instruments may be administered as often as necessary, they may be given less frequently than those used for instruction.
   - While the assessment tasks/items may represent those which are covered in the curriculum, the tasks/items are primarily indicators of possible success in particular areas of future study or employment.

INTRODUCTION

3. For use in program evaluation decisions:
   - Assessment may be evaluative with an emphasis on formulating policy or reviewing/developing the curriculum.
   - The assessment instrument may be administered as necessary but usually samples a broad range of performance indicators.
   - While the assessment tasks/items may represent those which are covered in the curriculum, many tasks may be too general to be useful in day-to-day decisions regarding a student's performance.
   - Although norm-referenced comparisons may be useful in examining an educational program, the norm group against which those comparisons are made should be carefully defined.

It is critical to decide on the specific uses for assessment/test results before developing or purchasing measures. If the purpose is to provide information for teacher use in making instructional decisions, the measures to be selected should focus on a narrow range of performance indicators. The measures should also be taken frequently, in order to incorporate new information on student learning into daily instructional decisions. However, if the purpose is related to educational policy or curriculum decisions rather than to instructional uses, the measures could cover a broad range of performance objectives and could be taken less frequently, possibly every two to four years. Likewise, assessment for guidance purposes may be infrequent and may cover a broad range of performance indicators.

(For more specific suggestions regarding the use of assessment/test results, see page 38.)

INTRODUCTION

WHAT OTHER IMPORTANT ISSUES RELATED TO TESTING MUST BE CONSIDERED?

- Who will coordinate the planning/implementation? Who will do the work necessary for implementation?
- What is the timeline for accomplishing each phase of implementation?
- What resources are necessary? What resources are/are not available locally?
- What is the cost of implementing the program? How will the cost be met?
- What evidence is there that the program as planned is fair to all regardless of color, national origin, race or sex? What assurances have been made that students will be held accountable only for skills which they have had an opportunity to learn?
- What evidence is there that the instruments used will yield scores which are valid and reliable?
- How will the program be evaluated?
- How will the competency assessment and testing program be coordinated with the existing standardized achievement/ ability testing program? With the ongoing evaluation of student performance? With the vocational education program offered to secondary school students?

INTRODUCTION

HOW CAN A DISTRICT GET STARTED?

- What is the district doing now? What should the district do by 1990?
- What should a management plan include? How will the plan be used?
- What approach could a district use to organize the implementation tasks? What are the advantages and disadvantages of each approach?

DECISION MAKING

Section 2
DECISION MAKING

The competency based education requirement states that pupil performance objectives will be established for English composition, mathematics and reading.

Appropriate assessment techniques will be used to assess student performance relative to the established objectives. Assessment techniques will include tests administered in at least three grades.

To be appropriate, the tests will be administered districtwide in the three grades and will measure student performance relative to the established objectives. The test results also must be used in a manner consistent with the locally established guidelines for use of test results. For example, a guideline might state that students may be held responsible only for those objectives taught.

Given these relationships, a district has a number of options from which to choose in answering the various questions posed in this section. Tables one, three, four and six are preceded by numbered options. The same numbers are used in the tables to refer to the defined options.

COMPETENCIES TO TEST

WHICH COMPETENCIES (PUPIL PERFORMANCE OBJECTIVES) SHOULD BE TESTED?

Each district is required to assess student performance relative to all objectives established in the competency based program. The following options, however, refer only to the districtwide tests, not to the ongoing assessment of pupil performance objectives.

OPTIONS:

1. Identify and test[1] only MINIMUM or critical skills at the specified grades.
2. Identify and test ALL important skills taught at the specified grades.
3. Identify and test skills which are REPRESENTATIVE of all those taught at the specified grades.
4. Select two or three of the above for use at different grades or in different subject areas.

[1]Districtwide tests of English composition are assumed to include an evaluation of the ability to compose in addition to the knowledge of principles or recognition of correct grammar. The English composition test may include one or more "prompts." The prompt serves the same purpose for written compositions that the item stem or question does for objective test items: it establishes the guidelines or provides a response set for the writer.

Table 1

Which Objectives to Test?
Advantages/Disadvantages of the Different Options

Option	Advantages	Disadvantages	Other Considerations
1	a. Fewer skills are tested. b. Intervention is needed for fewer students. c. The performance "floor" is established.	d. No inference can be made regarding the total program. e. The total instructional program may be narrowed.	f. Tests usually must be custom-made.

COMPETENCIES TO TEST

Table 1 *(Continued)*

Option	Advantages	Disadvantages	Other Considerations
2	a. Test information is available for the full range of the instructional program. b. Results may be used diagnostically to show student strengths/weaknesses.	c. Tests will be much longer than those in option one. d. The number of students identified for help may be large. e. Information will be difficult to manage without computer assistance.	f. Tests should be tailored to the local program. g. The ongoing competency based education program in all probability will include assessment of all identified skills.
3	a. Tests will be shorter than those in option two. b. Inference can be made about the full range of instructional program objectives.	c. The number of students identified as needing help may be large. d. Information about individuals may be difficult to manage without computer assistance.	e. Content sampling techniques will be of critical importance to the validity of the tests.
4	a. See the relevant items listed above.	b. See the relevant items listed above.	c. See the relevant items listed above.

ASSESSMENT TECHNIQUES

WHAT IS THE DISTINCTION BETWEEN ASSESSMENT AND TESTING?

ASSESSMENT as used in the competency standard is an umbrella term including any combination of the techniques defined in the list below. TESTING as used in the competency standard includes only those techniques beginning with item number 9 in Table 2. To meet the requirement for a TEST, the same instrument must be administered under similar conditions across an entire grade in the district; the items must measure performance objectives derived from the course of study; and the results must be used consistently with the guidelines established for using assessment/test results.

Table 2
Distinctions Between Assessment and Testing

Technique	Description
1. Observation/Anecdotal Record	A carefully prepared log of student behavior together with attending environmental characteristics.
2. Observation/Behavior Checklist	A simple checking of number of times a particular behavior occurred without regard for the circumstances.
3. Self Report Inventory	Students rate themselves on specified criteria.
4. Interview	A structured collection of student opinions in face-to face contact.
5. Rating Scales—Teacher or Peer	Usually a numerical scale with specified end points—e.g., Cooperative—Uncooperative. Completed by teacher or peers.
6. Ranking or Ordering by Merit	Peers or teachers rank a group of students on each of several traits.
7. Nominating or Other Sociometric Devices	Peers list people who fit a particular description.
8. Classroom Paper/Pencil Test	May be writing sample, completion, multiple choice or other format. Constructed by a teacher to measure objectives taught in his or her classroom.

ASSESSMENT TECHNIQUES

Table 2 *(Continued)*

Technique	Description
9. Locally Prepared Grade Level Test	May be writing sample, completion, multiple choice or other format. Measures objectives agreed upon by all teachers at the grade level (districtwide). Administered and scored uniformly in all classrooms.
10. Commercially Prepared Criterion-Referenced Test	Usually objectively scored, paper/pencil measures of specified pupil performance objectives. May be off-the-shelf tests or tailor made from large item banks. Results will always compare performance to the underlying criterion. Sometimes norms are also available.
11. Commercially Prepared Norm-Referenced Test	Usually machine scored multiple choice measures of curriculum objectives. Results will always compare performance to that of a specified norm group. Sometimes criterion-referenced information including item analysis is also available.

WHICH TEST TO USE

WHICH TEST SHOULD BE USED?

The options provided in this section refer specifically to the district-wide tests, not to classroom assessments.

When competencies (performance objectives) can be measured by selected-response tests (i.e., true-false, matching, or multiple choice), all of the following options apply. If students are required to demonstrate performance in other ways (e.g., write a composition, read aloud, or solve a mathematical problem using a calculator), only the fourth option will apply.

OPTIONS:

1. Select a standardized norm-referenced achievement test battery containing reading, language usage and mathematics items measuring the competencies identified by the district.
2. Select from commercial sources criterion-referenced tests measuring the competencies identified by the district.
3. Select from commercially available item banks the test items necessary to measure the competencies identified by the district.
4. Develop test items which will measure the competencies agreed upon by all teachers at a grade level, or devise "prompts" or problems which can be used to evaluate the identified categories of skills as in English composition or mathematics problem solving.
5. Use any combination of two or more of the above options.

WHICH TEST TO USE

Table 3
Advantages/Disadvantages for Test Options

Option	Advantages	Disadvantages	Other Considerations
1	a. The test is easily obtained. b. The test is relatively inexpensive. c. Items are reasonably well researched. d. Computer scoring and reporting are available.	e. The test may not suit the purpose for testing or skill area (e.g., does not measure minimums or writing abilities). f. Items may not be sensitive to indicate short term improvement after instruction. g. Competencies identified may be dictated somewhat by the test (e.g., the objectives covered in one part of the reading test may match well with local objectives; those in another part may not).	h. The relationship to the existing standardized testing program may be important. i. Inservice education is necessary for teachers on the use of results.

WHICH TEST TO USE

Table 3 *(Continued)*

Option	Advantages	Disadvantages	Other Considerations
2	a. The test is easily obtained. b. The test is relatively inexpensive. c. Computer scoring/reporting are sometimes available.	d. The test may be researched less well than the one in option one. e. The test may not suit the purpose. f. Competencies identified may be dictated somewhat by the test (e.g., the objectives covered in one part of the reading test may match well with local objectives; those in another part may not).	g. The relationship to the existing testing program and to informal measures (e.g., classroom tests which accompany the basal series) may be important. h. Inservice education is necessary for teachers on the use of results.
3	a. The test can be tailored to the local program. b. Computer scoring/reporting are sometimes available.	c. Test items are often not well researched. d. Tests are more expensive than in either of the previous options.	e. Relationship to existing testing program and to informal measures (e.g., classroom tests and tests which accompany the basal series) may be important. f. Inservice education for teachers is necessary on the construction of the tests and use of results.

WHICH TEST TO USE

Table 3 *(Continued)*

Option	Advantages	Disadvantages	Other Considerations
4	a. The test can be tailored precisely to local curriculum and instruction. b. Teachers in the district may improve their skills in constructing various measures of student performance. c. For English composition and certain mathematics problem solving skills, this type may be the only way to measure performance directly.	d. Computer scoring/reporting are often not available. e. This option requires enormous commitment in staff time.	f. Special efforts must be made to assure that the results will be valid and reliable. g. The relationship to existing testing program and to informal measures (e.g., tests which are used with basal series) may be important. h. Inservice education programs on constructing, scoring and interpreting the measures will be necessary.
5	a. See the relevant items listed above. b. The option may maximize efficiency and comprehensiveness.	c. See the relevant items above.	d. See the relevant items listed above. e. The assessment required for the ongoing competency based education program will in all probability combine several options.

CRITERIA FOR INSTRUMENTS

WHAT CRITERIA SHOULD BE USED IN SELECTING/CONSTRUCTING INSTRUMENTS?

Some assessment instruments including tests are available commercially. When deciding whether to use available instruments or to construct or select a new measure, the working committee should define what assessment or test information is needed. The following list of questions may be helpful.

- Who will use the assessment/test results?
- How will the assessment results be used in decision making? (For suggestions see page nine.)
- To whom and for what purpose will the information be interpreted or communicated?
- What is the most DIRECT measure of the specified skills or performance objectives?
- What type of score or feedback is necessary for the results to be useful?

After clearly defining the types of assessment results needed, the committee could use the following questions to direct its continuing search for the appropriate technique or instrument.

- How much of the necessary information can be obtained from the assessments and tests which are already used?
- Can the remaining information be obtained from instruments available commercially?
- Is it more efficient to design and produce measurement instruments locally than to purchase instruments tailored to the district's needs?
- Will the measures selected or constructed have high technical quality? (i.e., will they yield results which are valid and reliable?)
- Does the timeline favor (or perhaps even dictate) whether the instruments should be selected or constructed? (If the necessary instruments already exist, it usually takes less time to select than to construct a measure.)
- What other criteria are important in your district? (e.g., cost, time of administration, and norm or other standard for making comparisons)

For a more detailed treatment of the TEST SELECTION process, see the *Ohio Testing Handbook II* (Ohio Department of Education, anticipated publication date September 1983).

If a decision is made to CONSTRUCT the TESTS that will be used to satisfy the districtwide testing requirement, the following will be helpful:

CRITERIA FOR INSTRUMENTS

General guidelines for test construction.

- Standard directions for each task or group of tasks should be written and pretested.
- Time allotments, if appropriate, should be established and piloted in advance.
- Each measurement task should reflect accurately the pupil performance objective which it is designed to measure.
- Each measurement task should be clear-cut and concise.
- Tasks measuring the same objective should be independent, but similar in difficulty.
- Bias against subgroups (e.g., sex, ethnic) must be eliminated or controlled.
- Scoring methods and weights, if any, should be set in advance and communicated to the students taking the tests.

In developing appropriate ASSESSMENT techniques for measuring the pupil performance objectives established in the competency program, a district may use a combination of classroom tests, performance tests (including rating scales, checklists, anecdotal records, etc.), and teacher judgment. Guidelines for and selected examples of different assessment techniques are suggested below together with advantages and disadvantages for the two types of classroom achievement tests. Classroom achievement tests may be classified under two major headings: selected response types (e.g., multiple choice, true-false and matching) and constructed response types (e.g., completion, essay and problem solving).

SELECTED RESPONSE MEASURES have the following characteristics:

Advantages.

- Can be scored quickly and efficiently
- Can be scored objectively; that is, the correct answer is predetermined and any grader should arrive at the same test score
- If properly constructed, can measure higher level applications as well as simple recall
- Require less test-taking time for each question so more questions can be answered in a specified time period, and objectives can be sampled more extensively.

Disadvantages.

- Constructing accurate and precise items is difficult and time-consuming.
- Poorly constructed items often test recall of trivial knowledge.
- Correct answers can be achieved by guessing.
- Correct answers can sometimes be inferred from the item itself.

A test item that requires the student to CREATE a response rather than to select one is called a CONSTRUCTED RESPONSE item. It has the following characteristics:

CRITERIA FOR INSTRUMENTS

Advantages.

- Is comparatively easy to prepare
- Lends itself to measuring higher level cognitive objectives
- May yield additional insights to other learning

Disadvantages

- Is time consuming to score
- May be subject to the scorer's biases
- Offers only a limited coverage of a content domain (i.e., students respond to fewer questions)
- Often measures extraneous skills (e.g., knowledge of content may be measured in addition to prescribed writing skills)

If attainment of some specified objectives cannot be assessed by classroom paper and pencil tests, the teacher must devise another means of measuring achievement. PERFORMANCE MEASURES may be used in such instances. Such measures have the following characteristics:

- Each measurement task is subdivided into the various behavioral components necessary to complete the task satisfactorily.
- A checklist, rating scale or some other appropriate assessment instrument is adapted for use with the specified behavioral components.
- The teacher or other trained rater observes each student performing the tasks.
- The observer completes the appropriate instrument for each student. (For descriptions of appropriate techniques, see the list provided on page 16 in this section.)

For some pupil performance objectives, the most direct and practical assessment technique is TEACHER JUDGMENT. The question then is not "Should teacher judgment be used within the district's assessment plan?" but rather "How can teacher judgment make the overall assessment system efficient, comprehensive, and useful?"

Some conditions that will help "sharpen" the teacher judgment are

- Specifying clearly what learner behaviors should be observed
- Establishing in advance some decision rules (standards) for making judgments
- Observing the student performance under conditions which most likely will elicit the desired behavior
- Ensuring that the judgments are not affected unduly by irrelevant influences

Some irrelevant conditions or influences that have been identified (Oregon, 1977) are

- Public or teacher indifference about the characteristics to be judged
- Differences in interpretation of the competency characteristics that are to be judged. Friendships with students and/or parents

CRITERIA FOR INSTRUMENTS

- Apprehension over possible lack of administrative support
- "First impressions" of students
- Student appearances (unrelated to competency characteristics)
- Jumping to conclusions (based on unrelated observations)
- Guessing on the basis of inadequate or incomplete observation
- Prejudices or prior experiences
- Halo effects (coming from perceptions of general student abilities or attitudes)
- Comparing students to each other when this is inappropriate (e.g., when a criterion-referenced standard has been set)
- Rating students based on previous experience with other siblings
- Granting or withholding credit rather than asking for help in designing intervention activities that develop the desired characteristics in students
- Having a need for the class to "look good" (i.e., everyone acquired the characteristics to be judged)
- The standing of student's family in the community
- A lack of training and experience in the manner by which learners of various backgrounds may display desired characteristics
- Granting or withholding credit for competency attainment as a reward or punishment incentive, irrespective of student performance

For more detailed guidelines for constructing and using various types of tests and other assessment instruments, see the *Measurement Construction Guide* (Ohio Department of Education, anticipated publication date January 1984).

SPECIAL CONSIDERATIONS—ENGLISH

WHAT SPECIAL CONSIDERATIONS ARE NECESSARY FOR MEASURING ENGLISH COMPOSITION SKILLS?

At the district level, decisions about how to test student performance in English composition might be based upon what is known about the development of qualitative writing and about efficient and effective writing assessment. Brief statements on the latest research and illustrations of the assessment procedures follow. The research findings (Glatthorn, 1980) include:

- Teaching grammar is an ineffective way to teach writing. Composing a sentence requires mental activity that is different from that required for grammatical analysis.
- Writing is developmental. Sentence length, orderly word arrangement, and wider word choice increase with age and experience. Expectations for various age groups should be established accordingly.
- Emphasis should be placed upon the careful development (quality) of a limited number of papers with attention focused on solving communication problems before and during the writing, rather than on the hurried production of quantities of papers.
- Positive comments promote positive writing attitudes that result in improved writing skills.
- Instruction in mechanics is most effective within the context of developing and revising student writing samples.
- Writing activities designed to enhance orderly word arrangement will also lead to improved reading comprehension.
- Recently developed scoring procedures make it possible to reduce the time required for evaluating writing samples and also yield information that can be used for intervention decisions.

The following descriptions and comparisons of options may be useful in measuring student performance in English composition. Although there are many variations of methods, the following descriptions offer some generally agreed upon definitions.

OPTIONS:

1. Vocabulary tests—measure the understanding of and ability to see the relationships among words. (Sometimes used as placement or screening tools.)
2. Mechanics and usage tests—measure how well a person recognizes the language forms of edited American English, i.e., the English used by educated users of the language.

SPECIAL CONSIDERATIONS—ENGLISH

3. Essay Scale—involves the selection of a series of complete compositions arranged according to quality. At one end is an outstanding piece of writing, at the other end an inadequate one. The scale may range from four to eight categories. Raters attempt to match each student's writing with one of the models on the scale. This method was developed by the California Association of Teachers of English. More information about the Essay Scale is available from the National Council of Teachers of English.
4. Primary Trait Scale—attends to the mode of writing (e.g., narrative, descriptive, expository, or technical) and the features which are relevant to the kind of writing (e.g., audience, speaker role, purpose). The Primary Trait Scale may also be expanded to consider organization, word choice, and mechanics. This method was developed for the second writing assessment of the National Assessment of Educational Progress.
5. Analytical—involves an analysis of the writing that incorporates various categories of skills under two major headings: general merit (including ideas, organization, style, word choice, and sentence structure) and mechanics (including punctuation, capitalization, spelling, usage, and handwriting). Each composition is ranked by at least two trained evaluators using a graduated criterion-referenced scale. This option yields as many ratings (or scores) as there are categories of skills and may also yield an overall ranking. The score or ranking does not depend upon the mode of writing. Additional information is available from Educational Testing Service, Princeton, New Jersey.
6. Any combination of the above.

SPECIAL CONSIDERATIONS—ENGLISH

Table 4
Advantages/Disadvantages of the Methods for Testing English Composition Skills

Option	Advantages	Disadvantages	Other Considerations
1	a. Tests are hand/machine scored. b. Tests have high reliability.	c. Tests do not measure writing.	d. Tests may be used in a class or in larger groups.
2	a. Tests can be hand/machine scored. b. Tests provide efficient measures of how well a person recognizes commonly used language forms. c. Tests have high reliability.	d. High scores do not indicate ability to write.	e. Tests may be used in a class or in larger groups.
3	a. Scoring option yields an evaluation of writing as a whole. b. High reliability is possible if raters are trained.	c. The rating or score may not help identify specific strengths/weaknesses. d. This option assumes that a student's score is not affected by the mode of writing.	e. This rating method may be used with large as well as small groups of students (i.e., may be useful in the districtwide testing).

SPECIAL CONSIDERATIONS—ENGLISH

Table 4 *(Continued)*

Option	Advantages	Disadvantages	Other Considerations
4	a. High reliability is possible with training of raters. b. The method provides specific information on a narrowly defined writing task.	c. The student may receive a low score for good writing if a mode other than the one specified was used. d. This method usually requires more staff time than option three.	e. The amount of time required for scoring depends upon many factors including the number of features considered. f. Although certificated language arts teachers are recommended, others may be trained as raters.
5	a. With training, interrater reliability can be high. b. This method can be efficient in terms of rater time. c. The criteria for acceptable performance are set locally. d. Multiple ratings may facilitate intervention decisions.	e. Without proper training, rater may have difficulty maintaining consistent ratings from the first to the last paper.	f. This method assumes that a student's score is not affected by the mode of writing. g. The results may be useful for making instructional and intervention decisions.
6	a. Same as those listed above.	b. Same as those listed above.	c. Same as those listed above.

SPECIAL CONSIDERATIONS—ENGLISH

Some other considerations which apply across those methods that use student writing samples include:

- Papers should be scored anonymously.
- Raters should be trained to use a specific method—reliability and validity (to an extent) depend upon it.
- All students should respond to the same "prompts."
- Validity of the scores or ratings will depend heavily upon the procedures used to establish the rating scale(s).

In addition, the writing assessment should be fair to all students. The instruments should not be biased, nor should the results be used to discriminate against one or another group of students.

NOTE: For all methods requiring writing samples, students write a composition in response to a "prompt."

SPECIAL CONSIDERATIONS—READING

WHAT SPECIAL CONSIDERATIONS ARE NECESSARY FOR TESTING READING?

As plans are formulated for measuring reading competence, the following should be considered:

- In selecting or developing reading measures, consideration should be given to those reading objectives which cannot be measured in the usual paper and pencil "test." Some objectives identified in the overall competency education program may have to be addressed by direct observation of performance. The following are examples of such reading performance objectives. The student:

  reads words in a passage from left to right.

  expresses enjoyment when listening to prose or poetry read to the class.

  looks up unfamiliar words in the glossary.
- In selecting or developing reading comprehension tests, careful thought should be given to the stimulus passages. For example, should the passages be high interest, story type passages or informative, textbook-type passages? What difference will it make in inference about student performance?
- When selecting or developing reading tests, the readability of the material must be controlled carefully including the vocabulary and sentence structure.
- When selecting or developing reading tests, careful consideration should be given to assuring fairness (i.e., eliminating bias and discrimination).
- Before reading tests are administered, plans should include the establishment of an allotted time. If a test has a time limit (either explicitly defined as in a 40 minute standardized test or implicitly defined as in a test which is administered during one class period), then the speed of reading becomes a factor in a student's performance. To be untimed in the true sense, students may take as long as they need to finish every item. However, this is not always possible in every situation.
- As reading tests are selected or developed, consideration should be given to how the results will be used for intervention or instruction. For example, while criterion-referenced scores from various reading measures may directly affect day-to-day instructional decisions, norm-referenced scores (e.g., percentile ranks) may provide little information which is useful for making such decisions.

CONSIDERATIONS—MATHEMATICS

WHAT SPECIAL CONSIDERATIONS ARE NECESSARY FOR TESTING MATHEMATICS?

The following should be considered in preparation for measuring competence in mathematics:

- Care must be taken that measures of a student's ability to solve "word problems" or "story problems" do not become tests of reading instead of tests of mathematics.
- While some mathematics tests (e.g., tests of multiplication) may call for specific time limits, most tests requiring higher cognitive ability should be untimed as far as is reasonable.
- Problem contexts should be free of racial, cultural and sexual bias.
- Some mathematical objectives, for instance those concerned with making measurements or geometric constructions, may require actual performance of a skill, rather than reliance on inferential measures.
- If calculators or other electronic aids are used in development and reinforcement of certain mathematical skills or concepts, then consideration should be given to allowing their use on tests and assessments.
- If computer literacy or computer programming is part of the mathematics program, then it may be reasonable to include relevant computer objectives in the mathematics competency program.
- In a recent study of four widely used norm-referenced elementary mathematics tests, it was determined that each measured very different objectives at any given grade level (Freeman *et al.*, 1982). Test selection criteria, therefore, should include a match between the test and curriculum.
- "Knowledge" level items (i.e., multiplication facts, recall of definitions) are generally easier to test than objectives from higher cognitive levels (i.e., application, synthesis). Care should be taken that assessment and tests sample the entire range of cognitive level objectives taught and not focus only on lower level objectives.
- A focus on problem solving will necessitate flexible and creative testing and assessment.

ASSESSING MATHEMATICS

HOW CAN MATHEMATICS PROBLEM SOLVING BE ASSESSED/TESTED?

Since the ultimate goal of teaching school mathematics is to enable students to solve problems involving mathematics, problem solving is an important and appropriate component of a competency based education program. The testing and assessment of problem solving will require different approaches than open-ended or multiple choice paper/pencil tests. Measurement of problem-solving skills must be a continuous, long-term process conducted throughout the entire school mathematics program. Ideas offered in a previous discussion labeled WHAT USES CAN BE MADE OF ASSESSMENT/TEST RESULTS? (page nine) are especially important when considering the assessment/testing of problem solving. The following methods may be useful in assessing problem solving:

ATTITUDE MEASURES: The ability to solve problems is dependent in large measure on an individual's attitudes. The student must feel capable of obtaining a solution, must desire a solution and must believe that he or she can attack the problem. Therefore, assessing attitude is a vital component in making correct instructional decisions.

CHECKLISTS AND INVENTORIES: A checklist or inventory can be helpful in determining what a student knows about problem-solving strategies. Students may be given one or several problems and asked to solve each with a specified strategy (for example, draw a picture, make a graph, act out the problem, look for a pattern), or to solve a single problem in different ways using two or three specified strategies. In such situations it is more important to determine that the student can apply each strategy than to assure that the "correct" answer is given. Observations can be recorded on a checklist which will then serve as an inventory of skills.

INTERVIEWS: An interview may remove some limitations imposed by a written test. It is sometimes possible to examine more deeply how a student goes about solving a problem. A student's thinking can be inferred from the description of what he or she does. Basically, the interview situation should involve

- presentation of a problem
- the student's attempt to find a solution, describing what she or he is doing
- challenges or questions to the student, which should elicit specific details on what and why.

Notes should be made during the interview. If it is important to have exact records of student replies and comments, tape recorders may be used. Interviews may include one or more students.

ASSESSING MATHEMATICS

OBSERVATIONS: When students work individually or in small groups, the teacher should move about the room observing them and listening to them as they discuss problems, take notes, and ask questions. The goal of the observation is to gain information about how each student tries to solve problems. Some examples of points to consider include (Reys *et al.*, in press) the following:

- Do students read the problem carefully and accurately?
- Do students have a systematic way of attacking the problem?
- Do students apply an appropriate strategy, or do they use the last technique or topic that was taught?
- Do students try a different strategy if the first one fails?
- Do students apply a strategy in a consistent and persistent manner?
- Do students make careless errors? If so, when and what types?
- Do students concentrate on the task?
- Do students ask for help? If so, how quickly?
- Do students tend to attack all problems the same way?
- Do student behaviors and expressions indicate interest and involvement in the problem-solving task?

PAPER/PENCIL TESTS: Written tests should be one component of the assessment of problem-solving ability. Tests should follow the guidelines of the problem-solving program. Problems should be interesting and challenging. Students should have sufficient time to solve the problems. A valid measure of problem solving needs to contain new problems—NOT slight alterations of those previously studied. Many indicators of good problem solving can be measured by a paper/pencil test; however, problem-solving abilities can only be inferred. Certain subskills necessary for mathematical problem solving can be measured with a paper/pencil test, but these subskills may not be sufficient to make a good problem solver. Because problem solving has so many different facets, every effort should be made to utilize a variety of methods to measure performance.

WHEN TO TEST

WHEN SHOULD THE DISTRICTWIDE TESTS BE GIVEN?

This decision involves the specification of grade levels as well as the time of year for the testing. Illustrations of the many possibilities are included below.

OPTIONS for grade level

(Note: The standard requires districtwide testing in at least three grades.)

1. MINIMUM PLAN—may choose to test once in grades 1-4; once in grades 5-8; and once in grades 9-11.
2. ALTERNATE YEAR PLAN—beginning in either grade one or two, may test every other year through grade eleven.
3. ENTER/EXIT PLAN—may choose to test students as they enter and again as they exit a school. For example, if the school district is organized K-5, 6-8, and 9-12 then the students would be tested sometime during first grade, sometime during fifth grade, sometime during sixth grade, sometime during eighth grade, sometime during ninth grade, and sometime during eleventh grade.
4. COMPLETE ELEMENTARY/MINIMUM SECONDARY PLAN—may choose to test students every year during elementary years (i.e., 1-5, 1-6, or 1-8) and once during the senior high (9-11).
5. COMPREHENSIVE PLAN—may test every year, 1-11.

CONSIDERATIONS for determining grade level.

- Program Purposes—i.e., how the results will be used for instruction, evaluation, promotion, intervention, and guidance. (See the suggestions on pages 37-43 of this document.)
- District Organization—i.e., primary, elementary, middle and senior high or some other organizational pattern.
- Other Testing—i.e., what is done and at which grades.
- Human and Financial Resources—i.e., what will be the best combination to achieve program effectiveness in the most efficient manner? NOTE: While resources are required for testing and intervention, many of the existing resources may be used.

OPTIONS for time of year to administer district tests and advantages associated with each

1. Near the beginning of the year—results can be used immediately to effect classroom intervention and guidance decisions.
2. Near the end of the year—results can be used best for evaluation of curriculum, for various policy decisions, or for forming intervention groups at the beginning of the next year.

WHEN TO TEST

3. Throughout the year—tests need not cover all skills at once. Short tests focused on one or two skills may be given either before or after that skill is taught.
4. Pre-post test (either before and after teaching a particular skill or at the beginning/ending of a year or instructional unit) results can be used both for intervention and for evaluative purposes.

USING RESULTS

HOW WILL THE ASSESSMENT/TEST RESULTS BE USED?

The competency standard does NOT specify how assessment or test results will be used. Instead, each district is charged with establishing guidelines relating assessment/ test results to various decisions including instruction, intervention, promotion, evaluation, and guidance.

Table 5 below illustrates how guidelines on using assessment or test results can be applied to sample decision questions arising in each of the five categories. This list is NOT exhaustive.

USING RESULTS

Table 5
Sample Decisions and Guidelines
for the Use of Assessment Results

Decision Categories	Sample Decision Questions	Illustrations of Guidelines for Using Assessment Results
INSTRUCTION/ INTERVENTION	1. What is the best starting place for this individual or group? (Instruction)	1. The teacher will use the recommendations of other teachers, the student's performance on a locally constructed or selected pretest, and the subtest scores from last year's standardized achievement test to make such decisions.
	2. Which strategy is best for providing initial instruction and/ or intervention related to a particular sequence of skills?	2. Taking into account which strategies have been used successfully or unsuccessfully, the teacher will use professional judgment to select appropriate teaching strategies. (Assessment/test results may provide information about which strategies have worked—see the evaluation section of this table.)
	3. Is the student ready to go on to the next skill or unit of instruction (i.e., is remediation, reinforcement or support necessary)?	3. Alternative A: The teacher will use the results from appropriate assessment (including observation and judgment) to determine the student's readiness to progress. Alternative B: The student will perform each skill according to the prescribed standard (criterion) before going to the next skill.

USING RESULTS

Table 5 *(Continued)*

Decision Categories	Sample Decision Questions	Illustrations of Guidelines for Using Assessment Results
PROMOTION (Promotion decisions are usually based upon a series of instructional/ intervention decisions. NOTE: Graduation decisions are promotion decisions.)	4. Has the student demonstrated mastery of all the critical skills and shown satisfactory progress in learning other skills that were introduced at this level?	4. Alternative A: Mastery decisions will be based upon the criterion assessments made by the teacher. Satisfactory progress decisions will be based upon teacher observation and judgment. Alternative B: Students will demonstrate mastery of the critical skills by answering correctly a prescribed number of questions on the competency test. Satisfactory progress will be based upon teacher observation and judgment.
	5. Has the student acquired the skills necessary to succeed at the next higher level?	5. Such decisions will be based upon successful performance on the requisite skills test and upon teacher judgment. (Note: The accuracy of decisions is based in part upon a well articulated course of study.)
	6. Is the student ready to go on to the next higher level (i.e., the next course, post high school education, or a job)?	6. Readiness to proceed will be based upon successful completion of the work prescribed as judged by the teacher, mastery of the skills included in the critical skills test, and satisfactory completion of the entry level requirements for the next level as determined by teacher observation and judgment.

Table 5 *(Continued)*

Decision Categories	Sample Decision Questions	Illustrations of Guidelines for Using Assessment Results
GUIDANCE	7. Will the student in elementary school profit from an alternative educational experience (i.e., an accelerated mathematics program or a mentor program)?	7. Placement in alternative educational programs will be based upon teacher recommendations, the student's performance on the district's standardized achievement test, consent of the parent/guardian, and recommendation by specialists (e.g., school psychologist) when appropriate.
	8. Which high school courses or programs should the student elect?	8. The student will make such decisions after considering various school performance indicators (e.g., grades and achievement test scores in related areas) and results from career interest and/or aptitude measures.
	9. What kind of post high school education or training is appropriate for the student?	9. In making such decisions the student will consider high school grades; teacher recommendations; results of career interest inventories, vocational aptitude tests and college entrance examinations; and personal career goals.

USING RESULTS

Table 5 *(Continued)*

Decision Categories	Sample Decision Questions	Illustrations of Guidelines for Using Assessment Results
EVALUATION (PROGRAM EVALUATION)	10. Does the elementary school program prepare students adequately for instruction at the secondary level?	10. Such decisions will be based in part upon reports from elementary and secondary teachers and promotion rates at all grade levels.
	11. Does the current writing skills program prepare students as well for college entrance examinations as the previously used program?	11. Such decisions will be based upon longitudinal comparisons of results from the college entrance tests.
	12. Which instructional strategies seem to be the most effective?	12. The teacher, sometimes in consultaton with a supervisor or administrator, will make such decisions after carefully considering class and individual student performance, including grades in previous years, grades this year, achievement test scores, and competency test scores, in conjunction with teaching methods used.

Table 5 *(Continued)*

Decision Categories	Sample Decision Questions	Illustrations of Guidelines for Using Assessment Results
STUDENT EVALUATION	13. Is the student performing up to an expected level of competence?	13. Such a decision may be made by teachers, students and parents/guardians, in consultation with other professionals. The decision will be based upon the student's previous records in school, current classroom performance, and standardized test scores. The student's health, study habits and work schedule may also be considered.
	14. How will the student's final course grade be determined?	14. Alternative A: The student's course grade may reflect any combination of performance indicators deemed appropriate by the teacher in accordance with written policy. The procedure for determining the grade must be explained to the student and/or parent in writing. Alternative B: The student's final grade will be determined by the teacher using a combination of performance indicators, including teacher observation, classroom tests, homework, and the district's competency test. Procedures will be explained to the student and/or parent.

USING RESULTS

CONSIDERATIONS related to establishing guidelines for use of results.

- Input from the community, staff and students should be considered.
- The relationship between the guidelines for use of results and a number of other decisions discussed elsewhere should be considered (e.g., the type of assessment technique or test to use, how to set the standard for performance, time of year test is administered, length of the test needed, and the technical quality of the test items available).
- Consider the cost of implementing the guidelines for use of assessment results.
- Determine the probability of generating negative attitudes and decide how important this consideration is. Negative reaction will most likely occur if strong negative sanctions are associated with the use of results (e.g., when students must pass the test in order to be promoted).
- After deciding how the results will be used, consider the technical requirements of the instruments. For example, assessment devices used for promotion/graduation decisions will need to have validity established in ways that are different from those used for guidance purposes.
- After deciding how the results will be used, determine what type of information is needed from the assessment/test to support that decision. For example, while national percentile ranks or stanines may be helpful for some guidance uses or for identifying persons who may need special help, that information may not be useful in deciding whether the student has mastered a particular skill.

COMMUNICATING RESULTS

HOW WILL THE DISTRICTWIDE TEST RESULTS BE COMMUNICATED? TO WHOM?

ILLUSTRATIVE OPTIONS for reporting

1. Report individual results to teachers, students and parents. (This should be considered a minimum requirement whenever students are tested for whatever purpose.)
2. Report class summaries to the respective teachers.
3. Report building (grade level) summaries to all school staff.
4. Report district (grade level) summaries to all school staff.
5. Communicate building and district results to the residents of the district.
6. Record and retain individual test results for use at a later time.

CONSIDERATIONS for deciding what to communicate to whom.

- Program purpose—which results are reported to whom depends to a great extent on the intended use of the information.
- Family Privacy Act requirements regarding the release of individual student scores. For example, test information must be released to parents, but release of scores to anyone other than school staff may require the parent's permission.
- Professional Code of Ethics and Responsible Test Use as outlined by various professional groups including American Psychological Association, American Association for Counseling and Development (formerly called American Personnel and Guidance Association), American Educational Research Association, and National Council of Measurement and Evaluation.
- Legal implications for reporting or not reporting various test data.

SETTING STANDARDS

WHAT PERFORMANCE STANDARDS WILL BE SET?

Setting performance standards (i.e., defining acceptable and unacceptable performance) is highly subjective. Although there is no one objective method which yields the "right" answer, the following guidelines may help in making such decisions:

- Consider the standard setting process within the context of political, financial, legal and professional as well as technical issues.
- View the task as a complex, multiple step process.
- Base standard setting decisions upon all of the relevant data.
- Whenever possible, conduct a small scale tryout or pilot to see how the subjective standards work.
- Remember that the standards or criterion levels have NO meaning independent from the explicitly defined pupil performance objectives or the norm group used for a comparison.

Table 6 below defines and compares four general categories of standard setting procedures. It is assumed that the items/tests measure student performance relative to the established objectives.

OPTIONS for setting standards

(Criterion-referenced)

1. Identify a specific "cutting score" for student performance on each competency or skill measured by the test (e.g., three out of four correct, 80 percent correct, or a rank of five on a graduated scale for evaluating categories of skills in compositions).
2. Identify one specific "cutting score" for the entire test (e.g., 30 out of 40, 70 percent correct, or four on a five-point scale used in evaluating a composition).

(Norm-referenced)

3. Identify a national percentile rank (e.g., 25th) as the cut-off. Students who score below that point will receive remediation—much like the currently popular guidelines for screening students who will participate in various federally funded programs.
4. Using an acceptable statistical method for predicting expected achievement scores from academic ability test performance, identify students who are achieving significantly below their "expected performance."

NOTE: Any of the four preceding options may be used with standardized norm-referenced tests. The first three may be used with criterion-referenced tests (if norms are also available). Only the first two may be used with locally developed tests, including written compositions, unless local norms are developed.

SETTING STANDARDS

Table 6
Advantages/Disadvantages for Standard Setting Options

Option	Advantages	Disadvantages	Other Considerations
1	a. This method provides information which may be used for instructional decisions—relates to specific skills. b. Such standards are easy to set. c. This method is adaptable either to testing as each skill is taught or testing all skills at once in a single measure.	d. Such standards are usually arbitrary. e. The standards often do not reflect the fact that some skills are much more complex than others. f. Decisions are sometimes unreliable because of too few observations. g. This method is difficult to manage without computer assistance.	h. Reliability and validity of the decisions need to be established.
2	a. Such standards are easy to set. b. The information is easy to manage even without computer assistance. c. Decisions are usually more reliable because there are more observations.	d. Such standards are usually arbitrary. e. This method does not yield information which will help immediately in intervention or instructional decisions. f. This method usually takes no account of the complexity/difficulty of the measurement tasks.	g. This method is associated with a single test covering several objectives.

SETTING STANDARDS

NOTE: Although there are several ways to set a specific standard using either option one or two above, the comments usually apply regardless of method. For information on specific methods, see the appropriate part of the *Measurement Construction Guide.*

Option	Advantages	Disadvantages	Other Considerations
3	a. National percentile ranks are provided routinely when norm-referenced tests are used. b. This method provides a way to estimate the number of students who will need remediation. c. This method takes into account the complexity of the test items.	d. National percentile ranks provide little information which will help direct the intervention or instruction for those who do not meet the criterion.	e. Competency is defined in terms of group performance rather than in terms of what the individual can/cannot do.
4	a. This method does not require every student to achieve at the same level but emphasizes maximizing performance for everyone.	b. This method requires information about the student's academic ability. c. Computer support is necessary to establish the prediction equation. d. Little information is provided to help direct the intervention or instruction for those who do not achieve as predicted.	e. The method is based primarily upon statistical procedures, rather than observation or judgment.

PROVIDING INTERVENTION

WHAT TYPE(S) OF INTERVENTION WILL BE PROVIDED?

OPTIONS:

Many options for intervention are discussed and illustrated in another document prepared by the Ohio Department of Education. Some of these include:

1. Ongoing observation and diagnostic teaching
2. Teaching based upon the student's needs (e.g., learning style or modality)
3. Developmental learning and teaching
4. Alternative strategies including:
   - peer tutoring
   - cross-age tutoring
   - volunteers
   - grouping

CONSIDERATIONS

As intervention options are considered, some thought should be given to each of the following:

- Overall philosophy of the competency program
- Specific pupil performance objectives involved in each subject area at each grade
- Resources that are available, including who will do the intervention and what materials are needed
- The number of students who are involved
- Where and when the intervention will occur
- How long the intervention will take
- How the effects of the intervention will be demonstrated

HANDICAPPED

HOW WILL HANDICAPPED STUDENTS BE AFFECTED BY THE COMPETENCY PROGRAM?

Minimum standards require that written policies and procedures shall be adopted by the board of education regarding the participation of handicapped students in locally developed competency based education programs. Provisions may be made for exemption of individual handicapped students from the requirements of the locally developed competency based education programs.

The following questions and answers should be considered when developing these policies and procedures.

Question	Answer
1. Which students would be considered handicapped under these standards?	1. Rule 3301-51-01 U.- defines "Handicapped Child" as a person below twenty-two years of age who has one or more of the following handicaps: —Deaf —Deaf-Blind —Developmentally Handicapped —Hard of Hearing —Multihandicapped —Orthopedically Handicapped —Other Health Handicapped —Severe Behavior Handicapped —Specific Learning Disability —Speech Handicapped —Visually Handicapped
2. Can handicapped students be given a "blanket" exemption from the competency based education program?	2. Section 504 of the Vocational Rehabilitation Act of 1973 states that "no otherwise qualified handicapped individual... shall solely by reason of his/her handicap, be excluded from participation in, be denied the benefits of, or be subject to discrimination under any program or activity receiving federal financial assistance."

HANDICAPPED

Question	Answer
3. Can individual handicapped students be exempted from the competency based education program?	3. A handicapped student may be exempted on an individual basis. Standard 3301-35-02 (B)(2)(e) states that "provisions may be made for exemption of individual handicapped pupils from the requirements of the locally developed competency based education programs."
4. How may a handicapped student's participation in or exemption from competency based education programs be determined?	4. The following data contained on the student's Individualized Education Program should be examined when making this decision: —The statement of the present levels of educational performance of the student; —The statement of the specific special education and all related services to be provided to the student, and the extent to which the student will be able to participate in regular educational programs. Rule 3301-51-02 (E)(7)(a) (i) and (iii)

HANDICAPPED

Question	Answer
5. What specific modifications in assessment procedures must be made for handicapped students who participate in the competency based education program?	5. Rule 3301-51-02(D)(5)(c) states that "tests shall be selected and administered so as to best insure that when a test is administered to a child with impaired sensory, manual or speaking skills, the test results accurately reflect the child's aptitude or achievement level or whatever other factors the test purports to measure, rather than reflecting the child's impaired sensory, manual or speaking skills except wh[illegible] those skills are factors which the test purports to measure."

The following policy and procedures illustrate what may be used by school boards to meet this requirement. This sample policy statement has been judged to be in compliance with federal and state regulations regarding the education of handicapped students and is suggested as one method of meeting this requirement.

Sample Policy Statement

Handicapped students as defined in Ohio's Rules for the Education of Handicapped Children may be exempted from participation in or from the locally developed competency based education program requirements on an individual basis.

The determination that an individual handicapped student is exempted will be made using the following procedures:

- The student's present level of educational performance as indicated by the Individualized Education Program (IEP) will be reviewed.
- During the Individualized Education Program meeting, a statement regarding participation or exemption will be identified on the Individualized Education Program as part of the statement regarding the extent to which the child will be able to participate in regular education programs.
- If it is determined that the student will participate in the competency based education program, any necessary modifications in the testing procedure will be indicated on the Individualized Education Program.

OTHER ISSUES

Section 3

COMPETENCY TESTING: OTHER IMPORTANT ISSUES

In this section many of the important issues related to competency testing are identified and possible ways to deal with the issues are discussed.

TIMELINE

Part of the initial planning should consist of identifying what will be done, when it will be done and who will do it. Sample timelines are provided in Section Five of this document.

STAFFING

The success or failure of the competency based education program, just as with most other programs in the district, depends heavily on personnel. The following list provides an overview of several staffing needs:

- One person, responsible to the superintendent, should be appointed to coordinate the program districtwide.
- A steering committee involving personnel from each building and/or each grade level involved should be appointed to plan the overall direction.
- Subcommittees to oversee the work in each of the three content areas may be needed.
- As the need arises, task forces charged with specific short-term responsibilities should be appointed to complete the work.
- All staff, whether or not directly involved, should be kept informed during the entire project through periodic meetings, newsletters and inservice workshops.

RESOURCES

Planning should also include an analysis of the resources which may be needed to complete the project. Examples of such resources are listed below.

- Support materials — including tests, answer sheets, intervention materials, directions for test administration, guidelines for test use, student profile or record sheet, and newsletters.
- Computer hardware and software — for test scoring/ processing and perhaps for instructional management/record keeping.
- Facilities for test administration and for intervention.
- Human resources other than district staff, e.g., consultants to assist in test selection/development, in staff development, or in evaluation, as well as persons especially trained to score the writing samples.

OTHER ISSUES

Once a complete list of necessary resources is developed, an inventory of the district should be made. Are resources available locally? Could certain resources now being used for other programs be reallocated to this project for a specified time frame? After all available resources within the district are identified, what others are necessary?

COSTS

Cost projections should be developed as various options are considered. Although each district will have established procedures for estimating new program costs, the following suggestions may help:

- Divide the estimates into four phases:
  - Phase I — organization and planning costs
  - Phase II — development (including start up)
  - Phase III — ongoing implementation costs
  - Phase IV — evaluation

Implementation would include items such as purchasing and scoring tests and providing intervention materials and intervention personnel.

- Develop cost estimates for each option which is considered so that efficiency may be compared across methods.
- Identify whether or not each item represents a new cost or one that already exists (e.g., suppose that one option for measuring student competency is the standardized test currently used in your district. Costs associated with that test should be added into the overall projection, but identified in a way so that the coordinator could see that no new money is needed for purchasing tests).
- For the ongoing implementation phase, identify estimates in terms of per student cost. While the testing costs would be divided over all students tested, the intervention costs may be figured only in terms of those students involved in the intervention program. If possible, it would be desirable to estimate per student cost for each type of intervention.

ASSURING FAIRNESS

Confidence in a district's competency based education program is largely dependent upon fairness. Procedures should be established to assure that students will have equal opportunities to attain performance objectives. Instruction should be appropriate to individual needs. Judgment of students' educational needs should be fair, whether made during initial enrollment or throughout the learning continuum. Individual differences and other circumstances resulting from factors outside the individual's control should not work against the student. Decisions regarding instruction, guidance, intervention, promotion, and graduation should reflect fairness.

Fairness involves the absence of discrimination and bias. It is both illegal and unfair to provide educational programs which discrimi-

OTHER ISSUES

nate on the basis of color, national origin, race or sex. While it may be legal to use educational materials and activities which involve bias, discretion is recommended to assure fairness.

To assure that the competency based education program does not discriminate or involve bias, school administrators might consider the following:

- Assure that the daily expectations of educators, in addition to expectations implied by performance objectives, materials and activities, should challenge each student to perform according to his or her capacity regardless of color, race or sex.
- Establish assessment guidelines and procedures that are responsive to experiential factors and/or handicapping conditions which might impede or prevent a student from attaining performance objectives in the same manner or as rapidly as others. Among the experiential factors justifying intervention are
  - Previous experience (e.g., students transferring into the district might not have had appropriate opportunities to attain the objectives);
  - Language or dialect difference resulting from culture or socioeconomic circumstances.

  Handicapping conditions are addressed on pages 49-51.
- Establish a review panel including community representatives. Involve the panel in a review/advisory capacity for program materials, activities and policies.
- Establish guidelines to assure fairness in instruction, testing, decisions regarding students, and publication of test results. Students should be held accountable only for objectives taught.
- Review carefully the results of tryout or pilot studies to determine whether the tests seem to be fair for all groups.
- Communicate regularly with members of the staff and community regarding program purposes. Dissemination of pupil performance objectives and demonstrations of sample lessons and tests could enhance understanding and shared commitment to the program.

TECHNICAL CONSIDERATIONS

Whether the districtwide tests are purchased or developed locally, each district must assure the technical quality of the instrument(s) — e.g., the validity and reliability. Although the methods and documentations will vary depending upon the purpose of the program, the following suggestions may be helpful.

- Establish the relationship between each MEASURE and the COURSE OF STUDY.
- Establish that the item/test measures only the stated pupil performance objectives. For example, an arithmetic computa-

tion item should not measure a reading skill nor should a comprehension item require knowledge beyond what is given in the stimulus passage.

- Establish where and how each skill covered by the test is taught.
- Establish through tryouts that the item/test can assist in making "correct" decisions about students. (For example, that the standard is not set too high or too low.)
- Establish through tryouts that the item/test will measure performance consistently or that it will result consistently in the same decision regarding a student's need for intervention.

EVALUATION/REVIEW

Although the extent of the evaluation component will vary from district to district, every district must plan an evaluation/review. The following suggestions may be helpful.

- The review component could be based upon the management plan. (See page 61.) Procedures would address questions such as "Have we accomplished what we set out to do by a specified timeline?" and "How much did it cost?"
- The evaluation component should address questions regarding the IMPACT of the program on various groups or on specified targets.
- Evaluation targets may include among others:
  - student performance and attitudes
  - teacher performance and attitudes
  - curriculum
  - instructional materials
  - grading practices
  - promotion/retention policies
- Evaluation questions should be developed prior to the collection of data.
- Evaluation data may include impressionistic as well as "hard" data. Impressions may be gathered through surveys, structured interviews, questionnaires, etc. "Hard" data include measureable items such as standardized test scores, grades, attendance (teacher, student), scores on college entrance tests, promotion/retention ratios, and costs.
- Baseline data should be collected immediately and retained for comparison with similar data after the implementation is completed.

OTHER ISSUES

COORDINATION WITH EXISTING PROGRAMS

Guidelines should be established early in the planning phase to maximize efficiency, minimize conflict, and avoid duplication. The following areas are a few of those which should be considered.

- How will the competency testing program be coordinated with other testing in the district (e.g., with the standardized achievement/ability testing, the screening for compensatory programs, or the ongoing student evaluation)?
- How will the competency program affect students with identified handicaps? Although a district may include or exclude individual handicapped students from the program requirements, much thought should be given to the local implications before a policy is written. (For special considerations, see page 49 of this document.)
- How will the program be coordinated for secondary school students who attend a career center or joint vocational school? The need for coordination will be especially important to the joint vocational school which receives students from several districts each of which may have different requirements.
- How will the content of reading, English composition and mathematics be affected? For example, will the mathematics instruction be restricted only to the competencies tested?
- How will the content areas other than reading, English composition and mathematics be affected? For example, at the secondary level what responsibility will teachers of social studies, science, and other areas have for intervention or reinforcement of the identified skills?

GETTING STARTED

Section 4
GETTING STARTED

In 1983, not all districts are beginning at the same place. For some, the activities during the next few years will be analogous to redecorating or moving the furniture around in your house. For others the process may resemble adding a room. Still others will be remodeling the existing structure. Some may need to build a new structure "from the ground up." The information in preceding sections of this document will assist in deciding WHAT needs to be done. However, a BLUEPRINT (or management plan) will be necessary regardless of which decision path you follow. This section provides guidance for developing and implementing such a plan.

STUDY

CONDUCT A STUDY

Once a coordinator is assigned, the first task will be to conduct a study which will accomplish two tasks: determine where the district MUST BE or SHOULD BE by 1990 and identify where the district IS NOW. The study need not be lengthy, time consuming or costly.

To determine what the program should be by 1990, you may need to review data already collected and collect other information using several of the following methods:

- community survey
- analysis of test scores from last several years
- formal or informal meetings with teachers, students and parents
- review of research and literature on competency based education
- needs assessment
- conferences with representatives of other districts which have competency programs
- seminars with local planners including top administrators, representatives of teacher organizations and board members

To achieve the second objective (where the district is now) you will need to review the following:

- existing courses of study/curriculum guides
- existing pupil performance objectives
- existing guidelines and practices for using assessment and test results for instruction, intervention, evaluation, guidance, and promotion decisions (e.g., how are assessment/test results used in intervention decisions or in student evaluation/grading?)
- existing testing program as it relates to the course of study, to pupil performance objectives, and to classroom instruction (i.e., what relationship has been established among the curriculum, performance objectives, instruction and tests?)
- records of recently completed evaluations or reviews including those conducted internally and those conducted by persons from outside the district

The following checklist will help you determine where your implementation should begin, and help you locate information related to getting started. Because the steps are sequential, answer each question in turn. If the answer is "NO" or "?" (maybe, or I do not know), see the appropriate section referral.

STUDY

	Circle One
1. Are up-to-date, fully developed courses of study currently in place for English composition, mathematics and reading? If NO, see the publication entitled *Process Model for Course of Study* (Ohio Department of Education, 1983).	YES ? NO
2. Have pupil performance objectives (K-12) been established for English composition, mathematics and reading? If NO, see the appropriate section of *Process Model for Course of Study.*	YES ? NO
3. Has classroom instruction been adapted to the established pupil performance objectives? If NO, see the appropriate section of *Intervention* (Ohio Department of Education, 1983).	YES ? NO
4. Have assessment techniques been selected/developed to measure the stated pupil performance objectives in English composition, mathematics and reading? If NO, see Section Two of this document.	YES ? NO
5. Are tests (related to the previously identified course of study and performance objectives) administered uniformly across the district at least once in grades one through four, in grades five through eight, and in grades nine through eleven? If NO, see Section Two of this document for options.	YES ? NO
6. Have guidelines been established for the use of assessment and test results for instruction, intervention, evaluation, guidance and promotion decisions? If NO, see Sections One and Two of this document for suggestions and limitations.	YES ? NO
7. Is intervention (i.e., remediation, support and reinforcement) provided according to student needs? If NO, see *Intervention.*	YES ? NO

STUDY

	Circle One
8. Is there some provision to make another assessment <u>after</u> the intervention to determine whether or not student needs have been met? If NO, see Section Two of this document (Assessment Techniques).	YES ? NO
9. Is each component evaluated within five years of its completion date? If NO, see Section Three of this document for a list of possible evaluation questions and sources of data.	YES ? NO

Analyze the data related to both objectives to determine the gap or deficit between what is expected in the future and what exists now.

PLANNING

DEVELOP A MANAGEMENT PLAN (BLUEPRINT)

Once the data collection and analysis stages of the study have been completed, the results should show what needs to be accomplished before 1990. The MANAGEMENT PLAN has a beginning statement about WHAT EXISTS NOW, another regarding WHAT SHOULD EXIST BY 1990, and a final statement about WHAT NEEDS TO BE DONE IN THE INTERIM.

The management plan should also address each of the following:

- How the work will be coordinated? Provisions should be made for selecting an advisory or steering committee which will monitor the long range program activities.
- What work needs to be completed by what timeline? An outline of all tasks including a timeline for completion should be established.
- Who will do the work? Working committees should be established to develop each of the components identified in item two above. (How to staff these committees may also be addressed in the plan.)
- What resources will be required? Some thought should be given to whether or not outside personnel will be needed, whether or not additional instructional/intervention materials may be required, and whether or not a computer or microcomputer is necessary.
- What staff development/inservice opportunities will be required? Plans for inservice should include a discussion of appropriate topics, when to offer the sessions, and what consultants, if any, will be needed.
- What communication links need to be established? Consideration should be given to what to communicate, how, when and to whom.
- What baseline data need to be collected? Provisions should be made for collecting data now which can be used in later evaluations to determine the impact of the program.
- What will the program cost? Consideration should be given to resources which can be reallocated as well as to new financial resources which will be required.

ORGANIZATION MODEL

MODELS FOR ORGANIZING THE PROJECT

Once the study is completed and the necessary implementation tasks are identified, some organization must be attempted. For illustration purposes, assume that you decide to begin at STEP TWO (Establishing pupil performance objectives). There are at least two ways to organize the work load so all tasks will be completed by 1990. The following models are illustrations of these approaches.

1. SKILL AREA MODEL
   - One skill area is chosen (e.g., reading) and implemented K-12.
   - The second skill area is implemented (K-12).
   - The third skill area is implemented (K-12).

   NOTE: The sample timeline in Section Five illustrates this model.

   Advantage
   Provides an opportunity to articulate the skill strands (K-12), maximize efficiency, minimize undesirable redundancy, and assure that students have an opportunity to learn skills in a particular sequence.

   Disadvantage
   Unless there are overlapping timelines, one or more of the skill groups may have two years or less to complete all of the tasks. In addition, selection of this model usually precludes the use of a test booklet covering two or more course areas.

2. GRADE LEVEL ORGANIZATION MODEL
   - All skill areas (i.e., English composition, reading and mathematics) are implemented in grades K-4.
   - Once the work is well under way in the lower grades, implementation is begun in grades 5-8.
   - The same tasks are then implemented in grades 9-12.

   NOTE: The sample timeline provided in Section Five needs only slight modification to fit this model.

   Advantage
   Provides a systematic structure which could be used to assure that students have an opportunity to learn skills which will be taught and measured later.

   Disadvantage
   Unless there are overlapping timelines, some grade level units may be forced into very short time frames. In addition, unless careful articulation is emphasized across all grades, it is possible to develop for each skill area three unique, unrelated competency programs instead of one continuous K-12 program.

IMPLEMENTATION TIMELINE

Section 5
IMPLEMENTATION TIMELINE

In this section illustrative activities are identified and placed on a timeline to show how a district may proceed if it chooses. Developing a timeline for your district is one of the most important planning activities because it is helpful in program evaluation as well as in program implementation. To avoid disaster, implement the program components over an extended period of time. Do NOT adopt a full program during the first year.

Planning Year (1983-1984)

1. Appoint key leaders including the coordinator and key committees or task forces.
2. Review the school district's educational philosophy, course of study, instructional program and overall testing program, with particular emphasis on the relationships among these components.
3. Establish a competency program definition and philosophy for the district, including the specification of subject areas and the grade levels in which testing will occur and of how the test results will be used.
4. Identify persons in the district who have expertise in specific aspects of the program and who may be willing to help.
5. Establish a timeline for implementation including decisions regarding which course areas, grade levels, and tasks to initiate first, second, and third. NOTE: The decision regarding course area should be coordinated with the review schedule for the courses of study.
6. Collect baseline data which can be used later in an evaluation (e.g., attendance, test scores, promotion/retention ratios).
7. Hold seminars for staff, school board members and possibly community leaders to facilitate communication and planning, as well as to identify and develop expertise which may be needed later.
8. Initiate plans for determining lists of the competencies which will be measured. Establish the ground rules for all areas/grade levels; then begin more detailed planning in the one or two areas which are identified as the ones receiving first attention.
9. Establish guidelines for meeting the fairness requirements outlined on pages 53-54 of this document.
10. Establish guidelines for what is communicated to whom and when. Prepare and disseminate communications for the various target populations

IMPLEMENTATION TIMELINE

11. Identify the specific performance objectives (competencies) which will be assessed and tested in one subject area across all grade levels or in the specified grade range (e.g., K-4) across all three skill areas.

    NOTE: The following timeline illustrates the SKILL AREA MODEL described in Section 4. With only minor modifications, however, it could also reflect the GRADE LEVEL MODEL.

First Year of Implementation (1984-1985)

1. Decide which testing option will be used — i.e., — select existing instruments, construct measures, or both. (Note: This decision should NOT be made until AFTER the specific skill list is completed.)

2a. If TEST SELECTION, establish the test selection procedures. For assistance, see *The Ohio Testing Handbook II* (1983) available at cost from the Division of Educational Services, Guidance and Testing Section, in October 1983.

2b. If TEST CONSTRUCTION, decide who will be involved and plan staff development activities related to test construction. For assistance, see the *Measurement Construction Guide* (1984) available at cost from the Division of Educational Services, Guidance and Testing Section, after January 1984.

3. Involve the "Fairness Review Panel" and various task forces as the need arises.
4. Conduct supporting inservice workshops for either 2a or 2b.
5. Develop a plan for evaluating the procedures and the products.
6. Begin work on either test selection or test development for the first subject area or grade level group. Test selection will take several months so begin early and plan several sessions throughout the school year. If you are selecting tests, you may wish to allow time to try out some of the possible tests with a few students before the final decision is made. Test construction will usually take longer than test selection.
7. Communicate with all target populations.

Second Year of Implementation (1985-1986)

1. Develop alternate teaching strategies related to the skills being measured. In faculty meetings discuss procedures for integrating such strategies throughout the instructional process. For example, how can middle school social studies and science teachers use strategies which will strengthen or support these skills without weakening the social studies or science curriculum?

IMPLEMENTATION TIMELINE

2. Initiate plans for tracking a student's progress (across subjects and grades, as well as within one subject and grade).
3. Develop policies related to the handicapped student. (Note: While this decision may be made earlier, it may be helpful to wait until after some of the early decisions are made.)
4. Devise assessment instruments and procedures for the pupil performance objectives in the first selected course area or grade levels.
5. Administer three districtwide tests which were selected or constructed for the first subject area or grade level group. Try out all of the components including the three tests, the assessment procedures and the intervention strategies.
6. Evaluate the process and products.
7. Begin work on identifying pupil performance objectives in the second course area or at the second group of grade levels.
8. Continue communication with all appropriate groups.

Third Year of Implementation (1986-1987)

1a. If the instruments for the first test administration were selected, the pilot results may be used for determining:
- an effective set of standards
- an intervention rate that may be expected
- how well the test information is suited for the specific intervention purposes
- whether or not the tests appear to be nondiscriminatory (fair to all groups)

1b. If the tests were developed, the results may be used to determine the:
- technical adequacy of the items/tests
- areas of the tests which need improvement
- skills which are measured but should not be, or perhaps skills which should be measured but are not

2. Revise the assessments, tests, intervention strategies, or other components as needed.
3. Implement all revised components in the first course area or grade level group (full implementation).
4. Begin work on developing the components in the second area using different task force members, but the same steering committee if possible. Repeat the procedures which seemed to work earlier, but modify those procedures which did not work well.
5. Begin work on identifying pupil performance objectives in the third course area or in the third designated grade level group.
6. Evaluate the process and products.
7. Continue periodic communication with all groups.

IMPLEMENTATION TIMELINE

Fourth Year of Implementation (1987-1988)

1. Continue the assessment/testing and intervention in the first selected area.
2. Retest/reassess students to determine whether or not the intervention was successful.
3. Administer the three tests in the second selected course area or grade level group. Try out all components in this part of the program.
4. Revise the instruments and procedures in this second area.
5. Begin work on all components in the third area including assessment instruments, tests, and intervention strategies.
6. Plan how some of the skills in the first and second areas may be integrated into all subject areas. For example, if reading and English composition are the first and second areas introduced, teaching strategies associated with the tested skills should be integrated whenever possible into other instructional areas.
7. Complete development of the components in the third area including assessments, tests and intervention strategies.
8. Evaluate the process and products.
9. Continue to communicate.

Fifth Year of Implementation (1988-1989)

1. Continue assessment/testing/intervention/retesting in the first area.
2. Refine any of the components of the second area.
3. Implement all revised components in the second area, including intervention. (full implementation)
4. Retest students who participated in intervention in the second area.
5. Try out all components (including the administration of three tests) in the third area.
6. Evaluate the process and products.
7. Communicate periodically with all groups.

Final Year of Implementation (1989-1990)

1. Continue testing/intervention/retesting in the first and second areas.
2. Revise any components in the third area.
3. Implement all revised components in the third area. (full implementation)
4. Retest students who participated in intervention in the third area.
5. Evaluate the process and products.
6. Communicate periodically with all groups. (You may wish to prepare a "Status Report.")

IMPLEMENTATION TIMELINE

NOTE: Each of the "steps" identified in the cycles above will, of course, consist of many tasks. A district should define each of these steps as precisely as necessary and establish a separate timeline for each year.

Continue the established evaluation/review schedule after 1990.

BIBLIOGRAPHY

SELECTED BIBLIOGRAPHY

American Association of School Administrators. *Teaching Writing: Problems and Solutions.* Arlington, VA: AASA, 1982.

American Psychological Association. *Standards for Educational and Psychological Tests.* Washington, D.C.: APA, 1974.

Anastasi, A. *Psychological Testing* (5th edition). New York: MacMillan, 1982.

Association of English Teachers of Western Pennsylvania. *Suggestions for Evaluating Senior High School Writing.* Urbana, IL: NCTE, no date.

Berk, R.A. (Ed.). *Criterion Referenced Measurement: The State of the Art.* Baltimore, MD: Johns Hopkins University Press, 1980.

Brown, F.G. *Guidelines for Test Use: A Commentary on the Standards for Educational and Psychological Tests.* Washington, D.C.: NCME, 1980.

Buros, O.K. (Ed.) *The Eighth Mental Measurements Yearbook.* Highland Park, NJ: Gryphon Press, 1978.

California Association of Teachers of English. *A Scale for Evaluation of High School Student Essays.* Urbana, IL: NCTE 1960.

Clearinghouse for Applied Performance Testing. *Writing Assessment for the 80's.* Portland, OR: Northwest Regional Laboratory 1980.

Cooper, C.R. and Odell, L. *Evaluating Writing.* Urbana, IL: NCTE, 1977.

Cooper, C.R. (Ed.). *The Nature and Measurement of Competency in English.* Urbana, IL: NCTE, 1981.

Especially for Teachers: ERIC Documents on the Teaching of Writing, 1966-1981. Urbana, IL: ERIC Clearinghouse on Reading and Communication Skills, no date.

Fagan, W.T.; Cooper, C.R.; and Jensen, J.M. *Measures for Research and Evaluation in the English Language Arts.* Urbana, IL: ERIC/NCTE, 1975.

Freeman, D.J.; Kuhs, T.M.; Knappen, L.B.; and Porter, A.C. "A Closer Look at Standardized Tests," *Arithmetic Teacher.* 1982, 29(7), 50-54.

Glatthorn, A.A. *A Guide for Developing an English Curriculum for the Eighties.* Urbana, IL: NCTE, 1980.

Gronlund, N.E. *Measurement and Evaluation in Teaching* (3rd edition). New York: MacMillan Publishing Company, 1976.

Haley-James, S. (Ed.). *Perspectives on Writing in Grades 1-8.* Urbana, IL: NCTE, 1981.

Judine, Sister M., IHM. (Ed.). *A Guide for Evaluating Student Compositions.* Urbana, IL: NCTE, 1965.

Kasden, L.N. and Hober, D.R. *Basic Writing.* Essays for Teachers, Researchers, Administrators. Urbana, IL: NCTE, 1978.

Koch, C. and Brazil, M. *Strategies for Teaching the Composition Process.* Urbana, IL: NCTE, 1978.

Madaus, G.F. "NIE Clarification Hearing: the Negative Team's Case," *Phi Delta Kappan.* 1981, 63(2), 92-94. (Also, see Popham, 1981)

McCarthy, M.M. "The Applications of Competency Testing Mandates to Handicapped Children," *Harvard Educational Review.* 1983, 53, 146-164.

Mosenthal, P. "Defining Classroom Writing Competence: A Paradigmatic Perspective," *Review of Educational Research.* 1983, 53(2), 217-251.

Mosenthal, P.; Tamor, L.; and Walmsley, S.A. (Eds.). *Research on Writing: Principles and Methods.* New York: Longman's, 1983.

Mullis, I.V.S. Using the Primary Trait System for Evaluating Writing. Paper prepared for National Assessment of Educational Progress. Denver, CO: Education Commission of the States, 1980.

Mullis, I.V.S. and Mellon, J.C. Guidelines for Describing Three Aspects of Writing: Syntax, Cohesion and Mechanics. Paper prepared for National Assessment of Educational Progress. Denver, CO: Education Commission of the States, 1980.

Myers, M. *A Procedure for Writing Assessment and Holistic Scoring.* Urbana, IL: ERIC/NCTE, 1980.

Ohio Department of Education. *Competency Handbook.* Columbus, OH: ODE, 1978.

Ohio Department of Education. *Measurement Construction Guide.* Columbus, OH: ODE, in press.

Ohio Department of Education. *Ohio Testing Handbook: A Guide for Establishing an Effective School Testing Program.* Columbus, OH: ODE, 1977.

Ohio Department of Education. *Ohio Testing Handbook (II).* Columbus, OH: ODE, in press.

Oregon Department of Education. *Measuring Performance: Verifying Competencies Through Observation and Judgment.* Salem, OR: ODE, 1977.

Phi Delta Kappa. *Current Practices in the Implementation of Minimal Competency Testing Programs.* Cleveland, OH: PDK Greater Cleveland Inter University Chapter, 1981.

Phi Delta Kappa. *Descriptive Study of Minimal Competency Testing in Cuyahoga County—Research, Practices and Implications.* Cleveland, OH: PDK Greater Cleveland Inter University Chapter, 1978.

Popham, W.J. "The Case for Minimum Competency Testing," *Phi Delta Kappan.* 1981, 63(2), 89-91. (See also Madaus.)

BIBLIOGRAPHY

Popham, W.J. and Rankin, S.C. "Minimum Competency Tests Spur Instructional Improvement," *Phi Delta Kappan*. 1981, 62(9), 637-639.

Reys, R.; Suydam, M.; and Lindquist, M. *Helping Children Learn Mathematics*. New York: Prentice Hall, in press.

Roid, G.H. and Haladyna, T.M. *A Technology for Test-Item Writing*. New York: Academic Press, 1982.

Spady, W.G. "Competency Based Education: A Bandwagon in Search of a Definition," *Educational Researcher*. 1977, 6(1), 9-14.

Spady, W.G. and Mitchell, D. "Competency Based Education: Organizational Issues and Implications," *Educational Researcher*. 1977, 6(2), 9-15.

Stanford, Gene (Ed.). *How to Handle the Paper Load*. Urbana, IL: NCTE, 1979.

Sudol, R.A. (Ed.). *Revising*. New essays for teachers of writing. Urbana, IL: NCTE, 1982.

Thorndike, R.L. (Ed.). *Educational Measurement* (2nd edition). Washington, D.C.: American Council on Education, 1971.

Thorndike, R.L. and Hagen, E.P. *Measurement and Evaluation in Psychology and Education* (4th edition). New York: John Wiley and Sons, 1977.

Thurston, P. and House, E.R. "The NIE Adversary Hearing on Minimum Competency Testing," *Phi Delta Kappan*. 1981, 63(2), 87-88.

Trent, E.R. Summary of Results: 1983 Test Use Survey. Paper prepared for distribution to Ohio school superintendents. Columbus, OH: ODE, 1983.

Trent, E.R. and Golden, E. "Impact Evaluation of Locally Developed MCP's in Four Ohio Districts," *North Central Quarterly*. 1981, 56(1), 14-26.

NOTES

NOTES

NOTES

IMPLEMENTATION SERIES

A K-12 Written Guidance Plan

Code of Conduct

Competency Assessment and Testing

Economics Course of Study

Educational Options

Fine Arts and Physical Education K-12

High School Mathematics

Improving Attendance

Intervention

Materials Selection Guidelines

Procedures for School Evaluation

Process Model for Course of Study

School Health Services

Secondary Reading

For Additional Copies Write To:

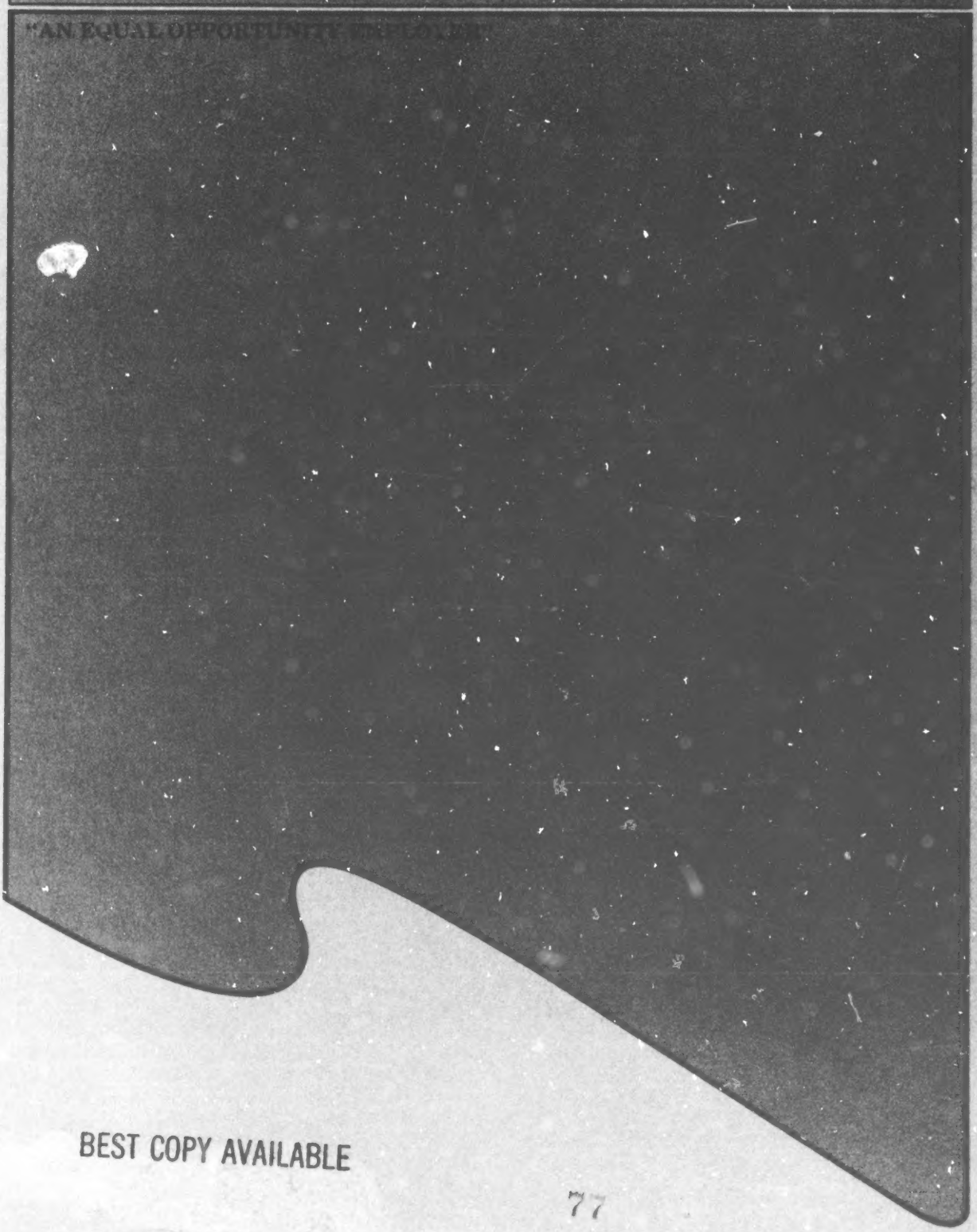
The Ohio Department of Education
Division of Elementary and Secondary Education
65 South Front Street
Columbus, Ohio 43215
614/466-2761

Total copies printed: 15,000
Unit cost: $.7296
Publication date: 9/83
(Includes paper costs)

BEST COPY AVAILABLE

THE OHIO DEPARTMENT OF EDUCATION
DIVISION OF ELEMENTARY & SECONDARY EDUCATION
65 SOUTH FRONT STREET
COLUMBUS, OHIO 43215 (614) 466-3741

"AN EQUAL OPPORTUNITY EMPLOYER"



BEST COPY AVAILABLE